

Heterozygous HMGB1 variants and neurodevelopmental disorder: phenotypic and allelic spectrum

Himanshu Goel

himanshu.goel@health.nsw.gov.au

University of Newcastle

Kayla Treat

Indiana University Health and Indiana University School of Medicine

Theodore Wilson

Indiana University Health and Indiana University School of Medicine

Angela Peron

Meyer Children's Hospital IRCCS

Manuela Morleo

Telethon Institute of Genetics and Medicine

Vincenzo Nigro

Telethon Institute of Genetics and Medicine

Konrad Platzer

University of Leipzig Medical Center

Rosanna Krakowsky

University of Leipzig Medical Center

Pongtawat Lertwilaiwittaya

University of Alabama at Birmingham

Bruce Keren

Pitié-Salpêtrière Hospital, AP-HP.Sorbonne Université, CHU Caen

Juan Pablo Trujillo Quintero

Parc Taulí Hospital Universitari, I3PT-CERCA, Universitat Autònoma de Barcelona

Nino Spataro

Parc Taulí Hospital Universitari, I3PT-CERCA, Universitat Autònoma de Barcelona

Flavio Faletra

ASU Friuli Centrale

Catia Mio

University of Udine

Christopher A. Walsh

Boston Children's Hospital

Jennifer Neil

Boston Children's Hospital

Henry Joel Mroczkowski

University of Tennessee Health Science Center

Sébastien Jacquemont

Centre Hospitalier Universitaire Sainte-Justine

Khadijé Jizi

Centre Hospitalier Universitaire Sainte-Justine

Deepika Burkardt

Children's National

Melissa Buckman

Tamworth Community Health Centre

Juliette Abrahamse

Prince of Wales Hospital

Article

Keywords: HMGB1, neurodevelopmental disorder, microcephaly, 13q12.3 microdeletion syndrome, BPTAS, nonsense-mediated decay

Posted Date: August 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10485683/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Heterozygous loss-of-function variants in *HMGB1* cause a neurodevelopmental disorder overlapping 13q12.3 microdeletion syndrome, whereas acidic-tail frameshifts cause brachyphalangy-polydactyly-tibial aplasia/hypoplasia syndrome (BPTAS). We sought to define the phenotypic and allelic spectrum and distinguish these disorders. Through international collaboration, we assembled 15 individuals from 13 unrelated families with heterozygous *HMGB1* variants and a neurodevelopmental phenotype. Phenotypes were harmonised using HPO terms. Variants were assessed for inheritance, NMD, splice effects, domain location and predicted C-terminal sequence. Developmental delay or intellectual disability was present in all individuals in whom development was assessed. Microcephaly occurred in 11/13 assessed and was occasionally severe. Recurrent findings included language impairment, hypotonia, behavioural difficulties, obesity, digital anomalies and craniofacial dysmorphism. Thirteen distinct variants comprised five predicted NMD alleles, three predicted NMD-escape alleles, three missense substitutions, one in-frame insertion and one de novo 5'UTR-proximal splice-region deletion. Two variants were transmitted by affected parents. The noncoding variant had moderate, concordant splice-loss predictions but lacked RNA confirmation. No C-terminal allele was predicted to create the arginine-rich basic tail characteristic of BPTAS. This cohort consolidates the *HMGB1*-related neurodevelopmental disorder that encompasses coding and noncoding alleles acting at transcript and protein levels. Microcephaly is common but not obligatory. Functional studies remain necessary for nontruncating and regulatory variants, whereas BPTAS is associated with a specific neomorphic basic-tail mechanism.

INTRODUCTION

HMGB1 (high mobility group box 1) is an abundant, highly conserved chromatin-associated protein that regulates DNA-templated processes. It contains tandem HMG-box DNA-binding domains (A and B) followed by an intrinsically disordered acidic C-terminal tail. (1–4) HMGB1 binds distorted DNA structures, bends DNA, modulates nucleosome accessibility and facilitates the recruitment of transcription and DNA-repair factors. (4, 5) Population constraint metrics indicate intolerance to loss of function, (6) and its constitutive loss of *Hmgb1* is lethal in mice, (7) consistent with an essential developmental role.

Three lines of evidence link *HMGB1* to human developmental disease. First, *HMGB1* lies within the minimal critical region of the 13q12.3 microdeletion syndrome and is considered a principal dosage-sensitive driver of its neurodevelopmental features. (8, 9) Second, Uguen et al. reported six individuals with heterozygous *HMGB1* loss-of-function variants or deletions a phenotype characterised by developmental and language delay, microcephaly, dysmorphism and variable obesity, establishing *HMGB1* haploinsufficiency as a cause of neurodevelopmental disorder. (10). Isolated *HMGB1* deletion cases have subsequently provided further support. (11) Third, *de novo* frameshift variants at the beginning of the acidic tail cause brachyphalangy-polydactyly-tibial aplasia/hypoplasia syndrome (BPTAS). These variants replace the acidic tail with an arginine-rich basic tail, alter biomolecular partitioning and produce nucleolar dysfunction through aberrant phase separation. (12)

The available evidence therefore suggests that *HMGB1*-related disease cannot be understood solely by variant position or by a coding-versus-noncoding distinction. Variants may reduce transcript dosage through nonsense-mediated decay (NMD), alter pre-mRNA processing, perturb the folded HMG-box DNA-binding domains, or create a new C-terminal protein with distinct biophysical properties. However, the neurodevelopmental disorder remains defined by few individuals, the contribution of nontruncating and regulatory alleles is uncertain, and transmission from affected parents has not been well documented.

Here we report 15 individuals from 13 unrelated families with heterozygous *HMGB1* variants and a neurodevelopmental phenotype. The cohort includes predicted NMD and NMD-escape alleles, missense substitutions in both HMG boxes, an in-frame insertion, and a de novo 5'UTR-proximal splice-region deletion. We sought to refine the phenotype, evaluate whether these mechanistically diverse alleles plausibly converge on reduced *HMGB1* output or activity, and clarify the molecular boundary between *HMGB1*-related neurodevelopmental disorder and BPTAS.

MATERIALS AND METHODS

Participants and ethics

Individuals were ascertained at international clinical genetics centres, with connections established through Genematcher. (13) Inclusion required a rare heterozygous *HMGB1* variant identified on clinical exome or genome sequencing together with a neurodevelopmental phenotype. Clinical and molecular data were collected retrospectively using a standardised proforma. The study was approved by the relevant institutional review boards, including HNE-2026-07-02_144640, with additional local approvals as required. Written informed consent was obtained for participation and, separately, for publication of clinical photographs. Consent forms specified that clinical photographs and de-identified clinical details might be freely available on the internet and seen by members of the public.

Clinical phenotyping

Participants were assessed by a clinical geneticist at the referring centres. Features were recorded as Human Phenotype Ontology terms (14) spanning growth, development, behaviour, neurological findings, dysmorphology and additional organ systems. Data included prenatal and growth history, developmental milestones, cognition and adaptive function, behaviour, neurological findings, neuroimaging, dysmorphology and other organ-system involvement. Growth parameters were expressed as standard deviation scores or centiles using locally appropriate reference standards. (15, 16) Facial photographs were reviewed centrally.

Molecular analysis and variant interpretation

Variants were identified by clinical exome or genome sequencing performed independently at each centre and confirmed by orthogonal sequencing. Segregation was assessed using parental samples where available. Variants are described on the MANE Select transcript NM_002128.7 according to HGVS nomenclature. (17) Classification followed ACMG/AMP principles, (18) informed by population frequency, (6) ClinVar (19) and available gene- and mechanism-specific evidence. Because several variants remain uncertain and functional testing was not uniformly available, classifications supplied by the contributing laboratories are reported without attempting a single retrospective reclassification.

Variant consequence prediction

Premature termination and frameshift variants were assessed for predicted NMD according to their position relative to the final exon-exon junction; variants in the terminal exon or within approximately 50–55 nucleotides of the final junction were considered likely to escape NMD. Predicted mutant coding sequences were checked using VariantValidator and Mutalyzer. (20, 21) For NMD-escape alleles, the predicted C-terminal peptide was examined for length, amino-acid composition and the arginine-rich basic tail characteristic of BPTAS, with charge estimates generated using ExPASy tools. (12, 22) Splice-region variants were evaluated with SpliceAI and Pangolin. (23, 24) Missense variants were evaluated using REVEL and AlphaMissense where available. (25, 26) In-silico results were treated as supportive rather than determinative, and splice outcomes without RNA evidence were explicitly designated as predictions.

Statistical analysis

Analyses were descriptive. Categorical findings are reported as n/N among assessed individuals. Given the cohort size, missingness and uneven ascertainment, genotype-phenotype observations were considered exploratory and no formal hypothesis testing was performed.

Facial gestalt and similarity analysis

Facial similarity was assessed using the GestaltMatcher framework. ((27, 28) For each image, GestaltMatcher generated Facial Phenotype Descriptors using a model ensemble with test-time augmentation, and pairwise distance in the Clinical Face Phenotype Space was used as the initial similarity metric, with smaller distances indicating greater similarity. For interpretability, pairwise distances were converted to a Gestalt Score, calculated as 1.3 minus the mean pairwise distance, so that higher scores indicate greater similarity. Cohort-level mean pairwise distances were calculated and converted to Gestalt Scores to build the HMGB1 intra-cohort distribution (100 values). A null reference distribution of 3540 values was generated by repeatedly sampling diagnosis-independent cohorts from the GestaltMatcher Database (28, 29) and converting their mean pairwise distances to Gestalt Scores in the same way. Each HMGB1 Gestalt Score was mapped to its empirical percentile within the random distribution using a right-inclusive empirical cumulative distribution function. As an image-level validation, eight HMGB1 facial images from seven individuals were compared against the GestaltMatcher gallery, and the gallery rank of each image was recorded when every other image was used as the query, with lower ranks indicating greater similarity; two of the eight images were longitudinal photographs of the same individual (Family 6, ages 5 and 9 years).

RESULTS

Cohort and molecular findings

Fifteen individuals from 13 unrelated families carried 13 heterozygous *HMGB1* variants (Table 1, Fig. 1). Five variants were predicted to undergo NMD: three nonsense variants, p.(Trp49*), p.(Gln142*) and p.(Tyr155*), and two frameshift variants, p.(Tyr16Cysfs*12) and p.(Glu40Valfs*3). Three variants were predicted to escape NMD and produce altered proteins: the terminal-exon frameshift p.(Glu215Lysfs*13); the splice-acceptor variant c.472-1G > A, predicted to generate a 13-residue C-terminal extension; and c.296 + 1G > A, predicted to introduce a frameshift and premature termination codon within 55 nucleotides of the final exon-exon junction, truncating the protein before the acidic tail. The latter two transcript outcomes require RNA confirmation. Three missense substitutions affected the folded DNA-binding domains: p.(Trp49Arg) in HMG box A, and p.(Ala149Pro) and p.(Lys154Asn) in HMG box B. One family carried the in-frame insertion p.(Glu47_Arg48insSerSerGlu) in HMG box A. Five coding variants clustered in HMG box A and four in HMG box B. Trp49 was independently affected by a nonsense and a

missense variant, and the adjacent residues Lys154 and Tyr155 were affected by a missense and a nonsense variant respectively, both ascertained at the same centre.

Table 1
Heterozygous *HMGB1* variants in the current cohort.

Family	cDNA (NM_002128.7)	Protein	Variant class	Predicted consequence	Region	Inheritance	Classification
F1	c.296 + 1G > A	p.?	Splice donor	Predicted frameshift/PTC within 55 nt of final junction; escape	Intron 3/exon 4	Maternal; de novo in mother	Likely pathogenic
F2	c.643_646del	p.(Glu215Lysfs*13)	Frameshift	Terminal-exon 13-residue C-terminal extension; escape	Acidic-tail terminus	De novo	Likely pathogenic
F3	c.424C > T	p.(Gln142*)	Nonsense	Predicted NMD	HMG box B	De novo	Likely pathogenic
F4	c.119_120del	p.(Glu40Valfs*3)	Frameshift	Predicted NMD	HMG box A	De novo	Likely pathogenic
F5	c.47_48del	p.(Tyr16Cysfs*12)	Frameshift	Predicted NMD	HMG box A	De novo	Likely pathogenic
F6	c.472-1G > A	p.?	Splice acceptor	Predicted 13-residue C-terminal extension; escape	C-terminal/HMG box B boundary	De novo	Pathogenic
F7	c.-14-14_-14-5del	p.?	5'UTR-proximal splice-region deletion	Possible acceptor weakening; reduced transcript output	5'UTR-proximal intron	De novo	VUS
F8	c.147G > A	p.(Trp49*)	Nonsense	Predicted NMD	HMG box A	De novo	Likely pathogenic
F9	c.445G > C	p.(Ala149Pro)	Missense	Substitution in folded DNA-binding domain	HMG box B	De novo	Likely pathogenic
F10	c.135_143dup	p.(Glu47_Arg48insSerSerGlu)	In-frame insertion	Three-residue insertion in DNA-binding domain	HMG box A	Paternally inherited	VUS
F11	c.145T > C	p.(Trp49Arg)	Missense	Substitution at same residue as F8	HMG box A	De novo	Likely pathogenic
F12	c.462A > C	p.(Lys154Asn)	Missense	Substitution adjacent to Tyr155	HMG box B	De novo (possibly mosaic)	Not yet classified
F13	c.465T > A	p.(Tyr155*)	Nonsense	Predicted NMD	HMG box B	Unknown; adopted	Likely pathogenic

Missense variant in HMG box B

An infant aged 7 months (Family 12) carried the de novo missense variant p.(Lys154Asn) in HMG box B, reported as possibly mosaic. This residue lies immediately adjacent to Tyr155, affected by the nonsense variant in Family 13, and both families were ascertained at the same centre. Developmental assessment was not yet possible at this age, and the variant has not been formally classified.

5'UTR-proximal splice-region variant

An additional 3-year-old female (Family 7) carried a *de novo* 10-bp deletion, NM_002128.7:c.-14-14_-14-5del (GRCh38 chr13:30463698 AAAAATAAAAT > A), identified by trio genome sequencing after nondiagnostic trio exome sequencing. The deletion lies in acceptor-proximal

intronic sequence immediately upstream of a 5'UTR exon. SpliceAI predicted acceptor loss (delta score 0.29 at -4 bp) with a secondary donor-loss signal (0.20 at + 463 bp); Pangolin predicted splice loss at the same local position (0.46 at -4 bp). No convincing cryptic splice gain was predicted. The possible consequences include inefficient 5'UTR splicing, intron retention, reduced normal transcript output or altered translation initiation. RNA studies were not available, and the variant remains a VUS.

Inheritance

Most variants were *de novo*. In Family 1, c.296 + 1G > A arose *de novo* in the affected mother and was transmitted to her son. In Family 10, the in-frame insertion was inherited from an affected father. The variant in Family 13 could not be tested because the individual was adopted and only proband sequencing was available. These two transmitting families demonstrate autosomal dominant inheritance compatible with reproduction and variable intrafamilial expression.

Growth, development and behaviour

Developmental delay or intellectual disability was present in all 14 individuals in whom development was assessed; the remaining participant was an infant aged 7 months at last review. Speech and language impairment was reported in 12 of 13 individuals for whom this domain was specifically assessed and was often disproportionate to other developmental abilities. Microcephaly was present in 11/13 assessed individuals and ranged from borderline to severe, reaching - 6.6 SD and - 4.22 SD in the most affected participants. It was absent in two individuals with nontruncating variants. Overweight or obesity occurred in several participants, but the spectrum also included failure to thrive, including Families 5, 6 and 7. Hypotonia was present in 6/11 assessed individuals. Behavioural features included autism spectrum disorder, attention-deficit/hyperactivity disorder, anxiety, aggression, self-injury, tics and depression, although ascertainment was incomplete.

Neurological and neuroimaging findings

Epilepsy occurred in three of ten assessed individuals, including the affected mother in Family 1. Brain MRI was available for seven individuals. Five were reported as normal. Family 7 had a small right parietal periventricular nodular heterotopia and a mildly dysmorphic corpus callosum, and Family 12 had microcephaly with a simplified gyral pattern, a small splenium of the corpus callosum, markedly diminished white matter and delayed myelination. These findings broaden the neurological spectrum but require replication before being considered characteristic of HMGB1-related disease.

Craniofacial and other features

The craniofacial phenotype was variable but recurrent elements included deep-set or short palpebral fissures, hypoplastic alae nasi, low insertion of the columella, thin upper lip, midface or malar hypoplasia and low-set or posteriorly rotated ears. For Family 6, serial photographs at four ages between 5 and 12 years document the evolution of the facial appearance, the triangular face, short downslanting palpebral fissures, hypoplastic alae nasi and pointed chin becoming more clearly recognisable with age (Fig. 2). These photographed ages also bracket a documented change in the clinical picture: an early assessment described developmental delay without frank intellectual disability, whereas formal testing at 12 years 7 months confirmed moderate intellectual disability with an intelligence quotient of 47, alongside new gross motor delay, tics and dental crowding. Over the same interval, occipitofrontal circumference moved from approximately - 5 SD to -3 SD, height from the 10th to the 10th-25th centile, and weight from below the 3rd centile to the 50th centile, so a persisting head-circumference deficit accompanied broadly typical somatic growth by adolescence. Whether the cognitive and motor findings reflect true evolution of the phenotype or increased detectability with age cannot be resolved from a single case, but the direction of change argues against anchoring prognostic counselling to an early assessment. Family 7 had hypotelorism, upslanting palpebral fissures, low insertion of the columella, a smooth philtrum, thin upper lip and low-set small ears. Digital or limb anomalies were documented in 8/10 assessed individuals and ranged from clinodactyly or brachydactyly to broad halluces, single palmar crease and, in Family 7, post-axial polydactyly. Other findings included hearing loss, refractive errors or strabismus, congenital heart disease, scoliosis, pectus excavatum, recurrent otitis media and severe eczema in one individual.

Major phenotypic findings in the current cohort are described in Table 2; the complete phenotype matrix for all 15 individuals is provided in Supplementary Table S1. Photographs of consenting participants are shown in Fig. 2

Table 2
Major phenotypic findings in the current cohort.

Feature	n/N assessed	Frequency	Comment
Developmental delay or intellectual disability	14/14	100%	Universal among those assessed; one infant not yet assessable
Speech or language delay	12/13	92%	Often disproportionate or severe
Microcephaly	11/13	85%	Common but not obligatory; occasionally severe (-6.6 SD)
Gross motor delay	7/8	88%	Ascertainment incomplete
Hypotonia	6/11	55%	Ascertainment incomplete
Epilepsy or seizures	3/10	30%	Variable age at onset
Digital or limb anomaly	8/10	80%	Usually minor; post-axial polydactyly in Family 7
Visual or ocular abnormality	5/6	83%	Refractive error, strabismus or microphthalmia
Feeding difficulty	5/7	71%	Includes gastrostomy in Family 4
Hearing loss	3/8	38%	Conductive or sensorineural descriptions varied
Overweight or obesity	4/10	40%	Failure to thrive in Families 5, 6 and 7
Abnormal brain MRI	2/7	29%	PVNH (Family 7); simplified gyral pattern (Family 12)

Facial appearance and intra-cohort similarity

Facial similarity across the cohort was quantified using GestaltMatcher, (27, 28) applied to eight facial images from seven individuals (Families 1 to 6 and 10). Compared with a random reference distribution, the HMGB1 Gestalt Score distribution showed a clear rightward shift, indicating greater intra-cohort facial similarity than expected by chance (Fig. 3). The HMGB1 cohort had a mean Gestalt Score of 0.4700 and a median of 0.4735, with individual scores ranging from 0.3476 to 0.6425. Mapping these values onto the random empirical cumulative distribution function placed the HMGB1 median at the 98.05th percentile of random; across all HMGB1 values, the percentile distribution had a median of 98.05%, a mean of 95.27% and a range of 58.33% to 99.94%. Image-level pairwise-rank analysis against the GestaltMatcher gallery supported this finding (Fig. 4); full rank matrix in Supplementary Table S3). The two longitudinal images of the Family 6 proband, taken at 5 and 9 years, correctly retrieved one another at rank 1 in both directions, confirming that the method recognised the same individual across a four-year interval. Strong cross-family matches were also observed, including reciprocal rank-1 matching between Families 1 and 4 and ranks 3 and 1 between Families 2 and 3. Together, these findings support a recognisable facial gestalt in HMGB1-related neurodevelopmental disorder.

Genotype-phenotype observations

The broad neurodevelopmental phenotype was shared across predicted transcript-reducing and nontruncating variant classes. Microcephaly appeared more frequent and severe among individuals with predicted NMD or NMD-escape alleles, whereas the two clearly nonmicrocephalic individuals carried a missense or in-frame allele. However, the nontruncating group was small, data were incomplete for several individuals and Family 13 had competing candidate variants; this pattern should therefore be regarded as hypothesis-generating. Family 7 adds a regulatory allele that may reduce HMGB1 output through abnormal RNA processing or translation, but functional confirmation is essential. None of the three NMD-escape variants was predicted to generate the arginine-rich basic tail characteristic of BPTAS.

DISCUSSION

This international cohort expands HMGB1-related neurodevelopmental disorder to 15 individuals from 13 families and addresses four unresolved questions: whether nontruncating variants can be associated with the phenotype, whether regulatory alleles may contribute, whether affected individuals can transmit the disorder, and how the neurodevelopmental spectrum is separated from BPTAS. The data support a consistent developmental disorder with frequent microcephaly and language impairment, but they also show that neither microcephaly nor obesity is obligatory.

The phenotype substantially overlaps that described by Uguen et al., (10) including developmental and language delay, microcephaly, dysmorphism and variable obesity. The present cohort broadens the spectrum to include more prominent behavioural morbidity, recurrent hearing and visual abnormalities, and frequent minor digital findings. The periventricular nodular heterotopia and mildly dysmorphic corpus callosum in Family 7 are notable because HMGB1 influences chromatin accessibility and developmental transcriptional programmes, but a

single observation cannot establish a recurrent neuronal-migration phenotype. The same caution applies to post-axial polydactyly in this individual, particularly because parental consanguinity raises the possibility of an undetected recessive or blended diagnosis despite nondiagnostic trio exome and genome analyses.

The allelic series is most informative when viewed across regulatory layers rather than as a simple collection of coding variants. Early truncating alleles are predicted to reduce transcript abundance through NMD and support haploinsufficiency. Missense and in-frame variants preserve the transcript but alter the folded HMG boxes, plausibly reducing DNA binding, DNA bending, nucleosome engagement or protein stability. Alleles that escape decay and extend or replace the C-terminus, such as the terminal-exon frameshift and the splice-acceptor variant, occupy an intermediate position: because they append residues beyond the normal termination site or remove the acidic tail, their consequence is governed less by nonsense-mediated decay than by ribosome-associated surveillance of aberrant elongation and termination, including non-stop and no-go decay, and by the stability of the altered product, so whether they lower or merely alter the protein is best resolved by RNA and protein assays rather than by prediction. The *de novo* c.-14-14_-14-5del variant is upstream of the coding start and may act through 5'UTR splicing, transcript abundance or translation efficiency rather than by directly changing an amino acid. Finally, acidic-tail frameshifts can create an abnormal protein with a new basic tail and altered phase-separation behaviour. *HMGB1* therefore provides a gene-specific example in which disease mechanisms can arise at the levels of RNA surveillance, RNA processing, translation, protein-domain function and intrinsically disordered-region biophysics. Across these levels the cohort alleles share a single direction of effect, complete loss of function in the null variants and hypomorphic loss in the missense and in-frame variants, with none predicted to be a gain of function; a toxic, neomorphic gain of function is confined to the basic-tail frameshifts of BPTAS.

This framework also defines the evidentiary limits of the current study. The coding truncating variants have a comparatively direct loss-of-function interpretation, but the missense, insertion and regulatory variants cannot be assumed to be pathogenic solely because they occur in *HMGB1*. The recurrence of distinct variants around Trp49, the localisation of nontruncating alleles within folded DNA-binding domains, *de novo* occurrence in several families and phenotypic overlap provide supportive evidence. Nevertheless, functional studies should test protein abundance, subcellular localisation, DNA-binding or DNA-bending activity, and phase-separation behaviour. For c.-14-14_-14-5del, RNA analysis should assess canonical junction usage, intron retention, cryptic acceptor selection, allele-specific expression and total *HMGB1* transcript abundance; protein quantification would be required if RNA structure is altered without a major change in abundance.

The distinction from BPTAS is mechanistically sharper than a simple positional boundary. In BPTAS, frameshifts beginning near residues Lys184 and Glu186 replace the acidic tail with a long arginine-rich basic tail, causing nucleolar mispartitioning and aberrant phase separation. (12) The C-terminal alleles in the present cohort do not reproduce that sequence transformation. This explains why an NMD-escape or C-terminal variant is not, by itself, sufficient to predict BPTAS. Conversely, minor digital anomalies, including post-axial polydactyly in Family 7, are not exclusive to the BPTAS mechanism. The predicted mutant peptide, particularly creation of the basic tail, is therefore more informative than the presence of any single limb feature.

Microcephaly was common and sometimes severe, but its absence in two individuals is clinically important. The predominance of postnatal onset suggests impaired brain growth rather than a primary cranial abnormality. *HMGB1* functions in chromatin organisation, transcription, DNA repair and cellular stress responses, all of which may influence neural progenitor proliferation, survival and differentiation. A future allelic series in neuronal or organoid models should compare a null allele, a DNA-binding-domain missense allele, a regulatory splice allele and a basic-tail BPTAS allele. Such a design could determine whether the neurodevelopmental alleles converge on reduced *HMGB1* dosage or activity while the BPTAS alleles produce a distinct neomorphic state.

Transmission from affected parents in two families has counselling implications. The disorder can be compatible with survival to reproductive age and transmission through either sex. For an affected heterozygous parent, the recurrence risk is 50% for each pregnancy, subject to the usual uncertainty regarding variable expressivity.

The study was retrospective, phenotyping and testing were not uniform, and ascertainment favoured clinically recognisable individuals. Several denominators are small because features were not systematically assessed. Longitudinal follow-up was available for only one family, so the direction and timing of the age-related change described for that individual cannot be generalised to the cohort as a whole. RNA or protein functional studies were unavailable for the splice, missense and insertion variants. Family 13 has competing *QRICH1* and *ELFN1* candidate variants, and Family 7 has parental consanguinity and a phenotype that could be blended; both should be interpreted cautiously. The apparent genotype-microcephaly relationship is exploratory. Objective facial analysis and genome-wide DNA methylation profiling may provide additional orthogonal evidence, although an *HMGB1* epismature has not been established.

In conclusion, *HMGB1*-related neurodevelopmental disorder encompasses a broader allelic and mechanistic spectrum than previously recognised. The core phenotype comprises developmental delay or intellectual disability with prominent language impairment, frequent microcephaly, variable behavioural and growth abnormalities, dysmorphism and minor limb anomalies. Coding loss-of-function alleles remain

the strongest evidence, while missense, in-frame and regulatory variants require functional confirmation. The boundary with BPTAS is best defined by creation of the arginine-rich basic tail and its neomorphic protein behaviour, rather than by C-terminal location alone.

Data Availability

The Human Phenotype Ontology-coded phenotype data and variant descriptions supporting this study are provided within the article, its tables and Supplementary Tables S1 to S3. Owing to the identifiable nature of clinical and photographic data, individual-level data are available from the corresponding author on reasonable request, subject to institutional and participant privacy constraints.

Declarations

Conflict of interest:

There is no conflict of interest

Ethical Approval

This study was approved by the relevant institutional review boards, including HNE-2026-07-02_144640, with additional local approvals obtained as required at each participating centre. Written informed consent for participation, and separately for publication of clinical photographs and details, was obtained from a parent or legal guardian for all participants under 18 years of age, in accordance with the Declaration of Helsinki.

Competing Interests

The authors declare no competing interests.

Funding:

No funding was received

Author Contribution

H.G. conceived and designed the study, prepared the figures and tables, and drafted the manuscript, H.G. and D.B. coordinated the international collaboration, H.G. and M.B. analysed the phenotype, H.G. and J.A. analysed the variants, K.T. and T.W.; A.P., M.M. and V.N.; K.P. and R.K.; P.L.; B.K.; J.P.T.Q. and N.S.; F.F. and C.M.; C.A.W. and J.N.; H.J.M.; and S.J. and K.J. ascertained and clinically characterised the individuals contributed by their respective centres, provided molecular findings and variant interpretation for those families, and reviewed the manuscript. All authors reviewed, edited and approved the final manuscript and agree to be accountable for its content.

Acknowledgments

The authors thank the participating families for their contribution to this study, and any facilitating clinicians, genetic counsellors or laboratories not listed as authors: to be completed by the authors.

Data Availability

All data supporting the findings of this study are available within the paper and its Supplementary Information.

References

1. Yamanaka S, Katayama E, Yoshioka K, Nagaki S, Yoshida M, Teraoka H. Nucleosome linker proteins HMGB1 and histone H1 differentially enhance DNA ligation reactions. *Biochem Biophys Res Commun*. 2002;292(1):268–73.
2. Yamada M, Ueda T, Sato K, Yoshida M. ATP-dependent chromatin structural modulation by multiprotein complex including HMGB1. *J Biochem*. 2004;135(1):149–53.
3. Lange SS, Vasquez KM. HMGB1: the jack-of-all-trades protein is a master DNA repair mechanic. *Mol Carcinog*. 2009;48(7):571–80.

4. Belgrano FS, de Abreu da Silva IC, Bastos de Oliveira FM, Fantappie MR, Mohana-Borges R. Role of the acidic tail of high mobility group protein B1 (HMGB1) in protein stability and DNA bending. *PLoS One*. 2013;8(11):e79572.
5. Blair RH, Horn AE, Pazhani Y, Grado L, Goodrich JA, Kugel JF. The HMGB1 C-Terminal Tail Regulates DNA Bending. *J Mol Biol*. 2016;428(20):4060–72.
6. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434–43.
7. Calogero S, Grassi F, Aguzzi A, Voigtlander T, Ferrier P, Ferrari S, Bianchi ME. The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. *Nat Genet*. 1999;22(3):276–80.
8. Bartholdi D, Stray-Pedersen A, Azzarello-Burri S, Kibaek M, Kirchhoff M, Oneda B, et al. A newly recognized 13q12.3 microdeletion syndrome characterized by intellectual disability, microcephaly, and eczema/atopic dermatitis encompassing the HMGB1 and KATNAL1 genes. *Am J Med Genet A*. 2014;164A(5):1277–83.
9. Mandrile G, Di Gregorio E, Calcia A, Brussino A, Grosso E, Savin E, et al. A new case of 13q12.2q13.1 microdeletion syndrome contributes to phenotype delineation. *Case Rep Genet*. 2014;2014:470830.
10. Uguen K, Krysiak K, Audebert-Bellanger S, Redon S, Benech C, Viora-Dupont E, et al. Heterozygous HMGB1 loss-of-function variants are associated with developmental delay and microcephaly. *Clin Genet*. 2021;100(4):386–95.
11. Wen T, Shayota BJ, Wallace L, Mani C, Davis N, Zhao J. A Case Report on 13q12.3 Microdeletion Syndrome Caused by HMGB1 Haploinsufficiency. *Case Rep Genet*. 2024;2024:1912620.
12. Mensah MA, Niskanen H, Magalhaes AP, Basu S, Kircher M, Sczakiel HL, et al. Aberrant phase separation and nucleolar dysfunction in rare genetic diseases. *Nature*. 2023;614(7948):564–71.
13. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015;36(10):928–30.
14. Robinson PN, Kohler S, Bauer S, Seelow D, Horn D, Mundlos S. The Human Phenotype Ontology: a tool for annotating and analyzing human hereditary disease. *Am J Hum Genet*. 2008;83(5):610–5.
15. Van den Broeck J, Willie D, Younger N. The World Health Organization child growth standards: expected implications for clinical and epidemiological research. *Eur J Pediatr*. 2009;168(2):247–51.
16. Nellhaus G. Head circumference from birth to eighteen years. Practical composite international and interracial graphs. *Pediatrics*. 1968;41(1):106–14.
17. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum Mutat*. 2016;37(6):564–9.
18. Tavtigian SV, Greenblatt MS, Harrison SM, Nussbaum RL, Prabhu SA, Boucher KM, et al. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genet Med*. 2018;20(9):1054–60.
19. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*. 2014;42(Database issue):D980-5.
20. Freeman PJ, Hart RK, Gretton LJ, Brookes AJ, Dalgleish R. VariantValidator: Accurate validation, mapping, and formatting of sequence variation descriptions. *Hum Mutat*. 2018;39(1):61–8.
21. Lefter M, Vis JK, Vermaat M, den Dunnen JT, Taschner PEM, Laros JFJ. Mutalyzer 2: next generation HGVS nomenclature checker. *Bioinformatics*. 2021;37(18):2811–7.
22. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res*. 2003;31(13):3784–8.
23. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019;176(3):535–48 e24.
24. de Sainte Agathe JM, Filser M, Isidor B, Besnard T, Gueguen P, Perrin A, et al. SpliceAI-visual: a free online tool to improve SpliceAI splicing variant interpretation. *Hum Genomics*. 2023;17(1):7.
25. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet*. 2016;99(4):877–85.
26. Cheng J, Novati G, Pan J, Bycroft C, Zemgulyte A, Applebaum T, et al. Accurate proteome-wide missense variant effect prediction with AlphaMissense. *Science*. 2023;381(6664):eadg7492.
27. Hsieh TC, Bar-Haim A, Moosa S, Ehmke N, Gripp KW, Pantel JT, et al. GestaltMatcher facilitates rare disease matching using facial phenotype descriptors. *Nat Genet*. 2022;54(3):349–57.
28. Lesmann H, Hustinx A, Moosa S, Klinkhammer H, Marchi E, Caro P, et al. GestaltMatcher Database - A global reference for facial phenotypic variability in rare human diseases. *Res Sq*. 2024.

Figures

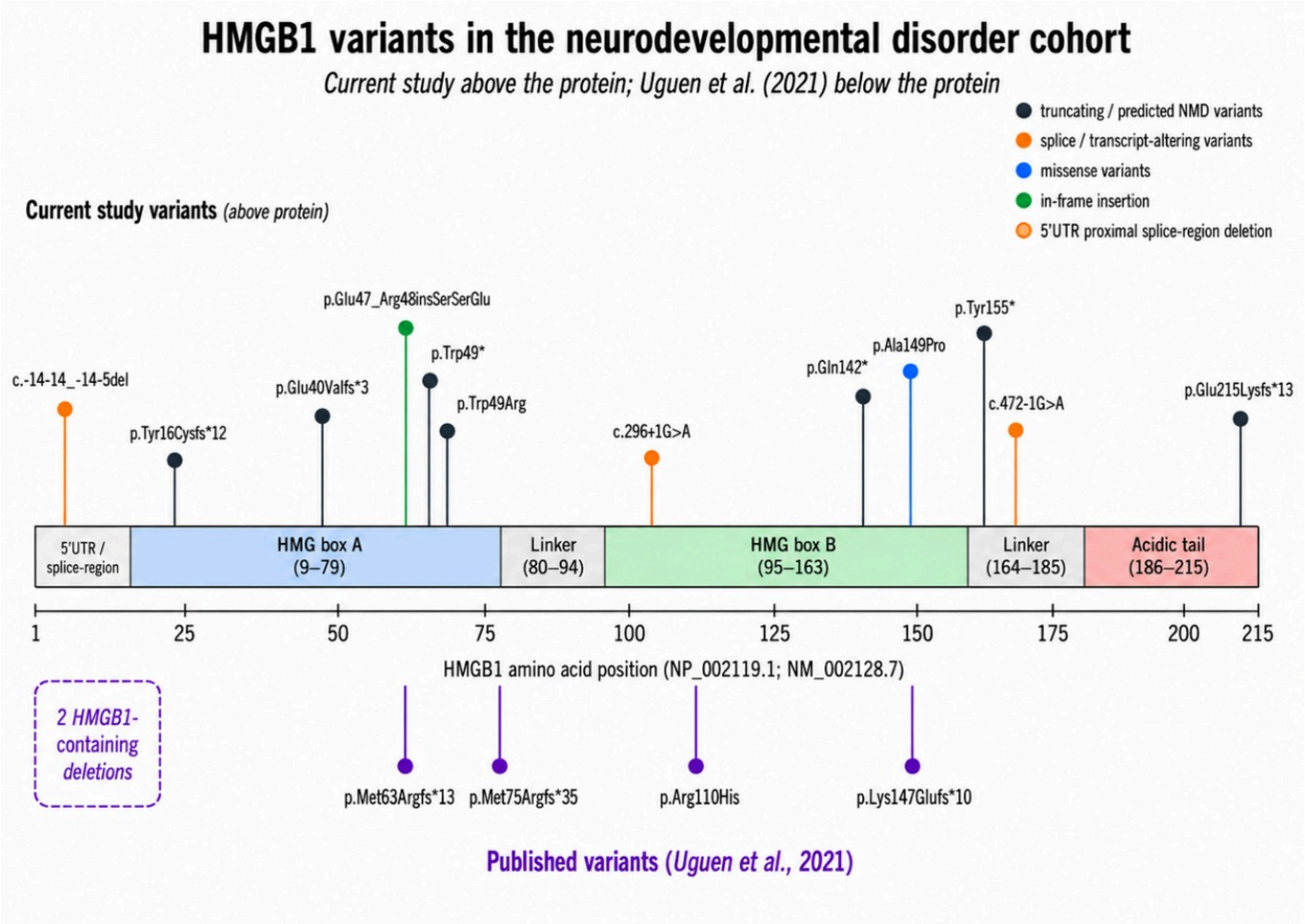


Figure 1

Distribution of *HMGB1* variants. Variants from the current cohort are shown above the 215-residue protein schematic; previously reported Uguen et al. sequence variants and *HMGB1*-containing deletions are shown below, together with published BPTAS acidic-tail frameshift variants. The c.-14-14_-14-5del variant is shown separately because it lies upstream of the coding start. HMG boxes A and B and the acidic tail are indicated.

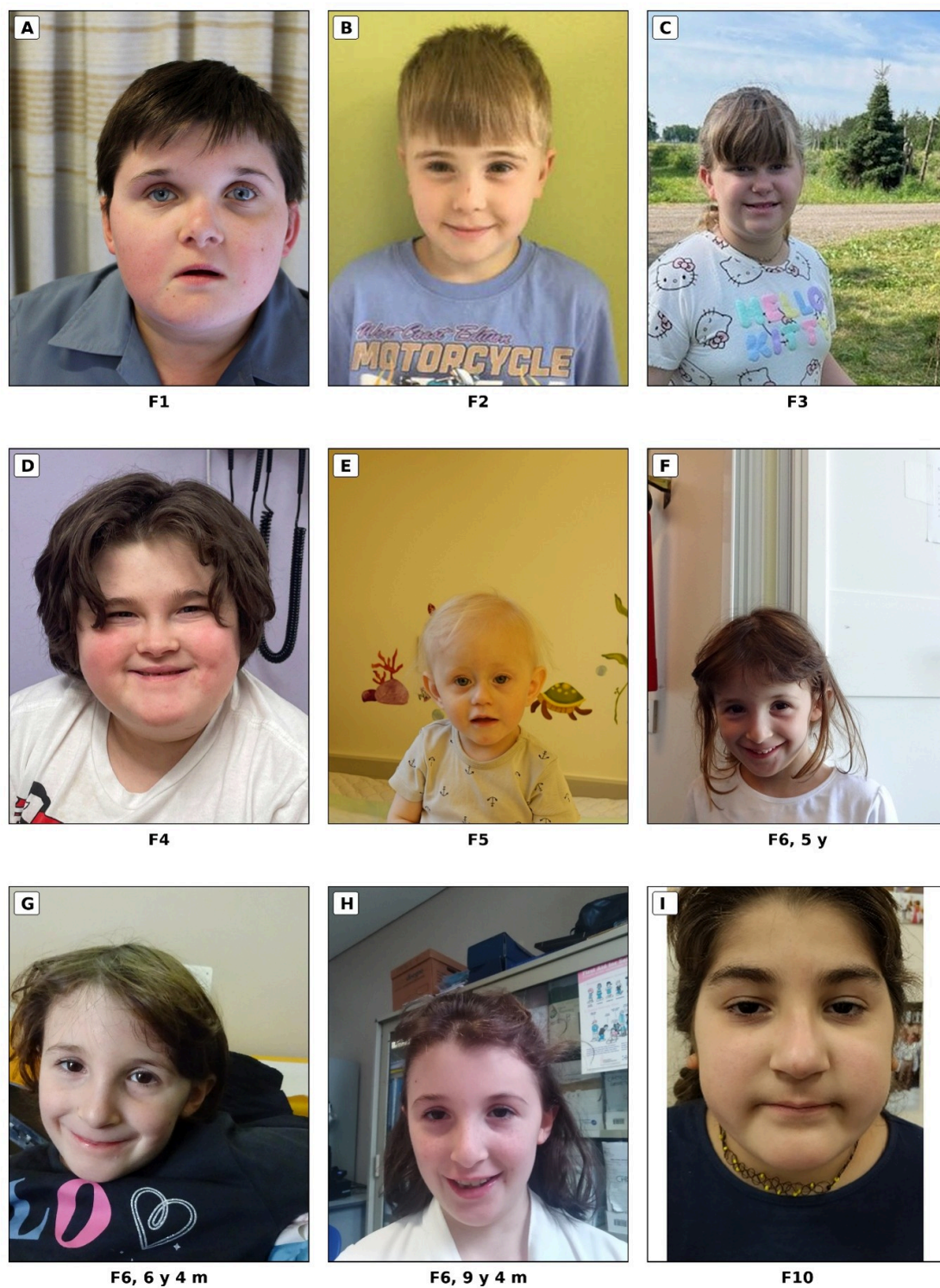


Figure 2

Facial appearance of individuals with HMGB1-related neurodevelopmental disorder. Images are shown only for participants with specific publication consent. A longitudinal series for Family 6 illustrates age-related evolution of the facial appearance.

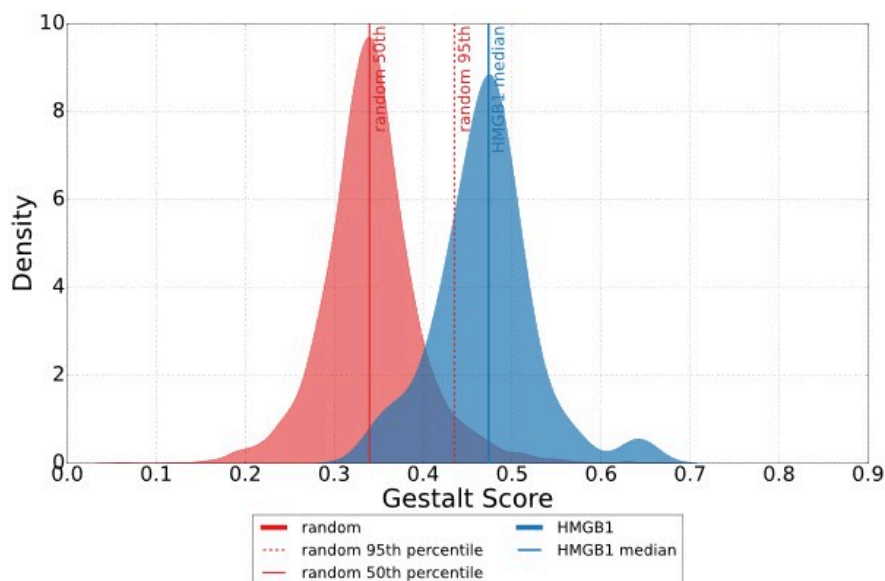


Figure 3

HMGB1 versus random intra-cohort Gestalt Score distributions. Kernel density estimates of Gestalt Scores for the HMGB1 cohort (blue) and a random reference distribution (red). Higher Gestalt Scores indicate greater facial similarity. Vertical lines mark the random 95th percentile (0.4351), the random median (0.3397) and the HMGB1 median (0.4735). HMGB1 Gestalt Scores had a mean of 0.4700, a range of 0.3476 to 0.6425 and an interquartile range of 0.4352 to 0.4949. The rightward shift indicates greater intra-cohort facial similarity than expected under random sampling.

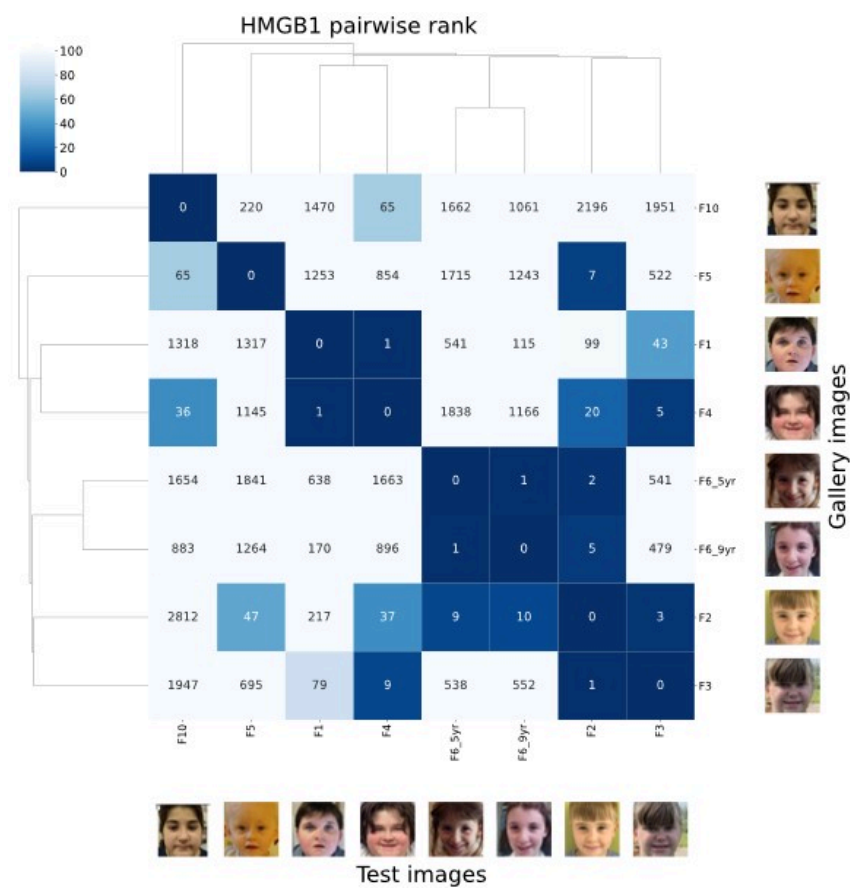


Figure 4

HMGB1 pairwise ranks in the GestaltMatcher gallery. The matrix is organised column-wise: each column represents an HMGB1 query image and each row an HMGB1 candidate image. Cell values indicate the gallery rank of the row image when the column image was queried, with

smaller ranks indicating greater similarity. The longitudinal images of the Family 6 proband at 5 and 9 years retrieved one another at rank 1 in both directions. Additional low-rank matches, including between Families 1 and 4 and between Families 2 and 3, highlight localised facial similarity within the cohort.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)
- [SupplementaryTablesS1S3.xlsx](#)